MATHEMATICAL SURVEYS
NUMBER I

THE PROBLEM OF MOMENTS

BY
J. A. SHOHAT
UNIVERSITY OF PENNSYLVANIA
AND
J. D. TAMARKIN
BROWN UNIVERSITY

PREFACE

The problem of moments is of fairly old origin, but it received its first systematic treatment in the works of Tchebycheff, Markoff, Stieltjes, and, later, Hamburger, Nevanlinna, M. Riesz, Hausdorff, Carleman, and Stone. The subject has an extensive literature, but has not been treated in book or monograph form. In view of the considerable mathematical (and also practical) interest of the moment problem it appeared to the authors desirable to submit such a treatment to a wide mathematical public. In the present monograph the main attention is given to the classical moment problem, and, with the exception of a few remarks concerning the trigonometrical moment problem, no mention is made of various generalizations and modifications, important as they may be. Furthermore, lack of space did not permit the treatment of important developments of Carleman and Stone based on the theory of singular integral equations and operators in Hilbert space. On the other hand, a special chapter is devoted to the theory of approximate (mechanical) quadratures, which is intimately related to the theory of moments and in many instances throws additional light on the situation.

The bibliography at the end of the book makes no claim to completeness.

The authors wish to acknowledge with gratitude the help received from Brown University and the University of Pennsylvania (through its Faculty Research Committee) in preparing this manuscript for publication.

INTRODUCTION

1. Brief historical review. In 1894–95 Stieltjes published a classical paper: "Recherches sur les fractions continues" containing a wealth of new ideas; among others, a new concept of integral—our modern "Stieltjes Integral". In this paper he proposes and solves completely the following problem which he calls "Problem of Moments":

Find a bounded non-decreasing function $\psi(x)$ in the interval $[0, \infty)$ such that its "moments" $\int_0^\infty x^n \, d\psi(x)$, $n = 0, 1, 2, \cdots$, have a prescribed set of values

$$\int_0^\infty x^n \, d\psi(x) = \mu_n , \qquad n = 0, 1, 2, \cdots . \tag{1}$$

The terminology "Problem of Moments" is taken by Stieltjes from Mechanics. [Stieltjes uses on many occasions mechanical concepts of mass, stability, etc., in solving analytical problems.] If we consider $d\psi(x)$ as a mass distributed over $[x, x + dx]$ so that $\int_0^x d\psi(t)$ represents the mass distributed over the segment $[0, x]$—whence the modern designation of $\psi(x)$ as "distribution function"—then $\int_0^\infty x \, d\psi(x)$, $\int_0^\infty x^2 \, d\psi(x)$ represent, respectively, the first (statical) moment and the second moment (moment of inertia) with respect to 0 of the total mass $\int_0^\infty d\psi(x)$ distributed over the real semi-axis $[0, \infty)$. Generalizing, Stieltjes calls $\int_0^\infty x^n \, d\psi(x)$ the n-th moment, with respect to 0, of the given mass distribution characterized by the function $\psi(x)$.

Stieltjes makes the solution of the Moment-Problem (1) dependent upon the nature of the continued fraction "corresponding" to the integral

$$\begin{aligned} I(z, \psi) = \int_0^\infty \frac{d\psi(y)}{z + y} &\sim \frac{\mu_0}{z} - \frac{\mu_1}{z^2} + \frac{\mu_2}{z^3} - \frac{\mu_3}{z^4} + \cdots \\ &\sim \frac{1|}{|a_1 z} + \frac{1|}{|a_2} + \frac{1|}{|a_3 z} + \frac{1|}{|a_4} + \cdots, \end{aligned} \tag{2}$$

and upon the closely related "associated" continued fraction

$$\frac{\lambda_1 |}{|z + c_1} - \frac{\lambda_2 |}{|z + c_2} - \frac{\lambda_3 |}{|z + c_3} - \cdots \tag{3}$$

derived from (2) by "contraction":

$$z - \frac{\alpha|}{|1} - \frac{\beta|}{|z - \gamma} = z - \alpha - \frac{\alpha\beta}{z - (\beta + \gamma)} .$$

Making use of the theory of continued fractions Stieltjes shows that in (2) all a_i are positive (which results in the positiveness of all λ_i and c_i in (3)).*

He further shows that this necessary condition is also sufficient for the existence of a solution of the Problem of Moments (1). In terms of the given sequence $\{\mu_n\}$ this condition is equivalent to the positiveness of the following determinants

$$(4)\qquad \Delta_n = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix} \equiv |\mu_{i+j}|_{i,j=0}^{n}; \qquad n = 0, 1, 2, \cdots,$$

$$(5)\qquad \Delta_n^{(1)} = \begin{vmatrix} \mu_1 & \mu_2 & \cdots & \mu_n & \mu_{n+1} \\ \mu_2 & \mu_3 & \cdots & \mu_{n+1} & \mu_{n+2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mu_{n+1} & \mu_{n+2} & \cdots & \mu_{2n} & \mu_{2n+1} \end{vmatrix} \equiv |\mu_{i+j+1}|_{i,j=0}^{n}; \qquad n = 0, 1, 2, \cdots.$$

The solution may be unique, in which case we speak of a "determined Moment-Problem"; or there may be more than one solution in which case there are, of necessity, infinitely many solutions; our Moment-Problem is then "indeterminate". Stieltjes illustrates the latter case by a remarkable example. He further gives an effective construction of certain solutions of the Moment-Problem (all, of course, essentially the same in case of a determined problem) which in the indeterminate case turn out to possess important minimal properties. Here the denominators of the successive approximants to the continued fractions (2) and (3) play an important role. In passing Stieltjes introduces an important new proposition dealing with the convergence of series of functions of a complex variable (now known as the Stieltjes-Vitali Theorem) which leads to a complete solution of the problem of convergence of the continued fraction (2) in the complex z-plane. Here Stieltjes shows that the Moment-Problem (1) is determined or indeterminate according as the continued fraction (2) is convergent or divergent, that is, according as the series $\sum_1^\infty a_i$ diverges or converges. The interesting fact that the continued fraction (2) may converge for certain z (to the value $I(z, \psi)$), while the series $\sum_0^\infty (-1)^i \mu_i z^{-i-1}$ diverges for all z is demonstrated.

* In the subsequent discussion we write

$$\int_0^\infty \frac{d\psi(y)}{z-y} \sim \frac{\mu_0}{z} + \frac{\mu_1}{z^2} + \frac{\mu_2}{z^3} + \cdots,$$

so that the corresponding and associated continued fractions (2) and (3) are replaced respectively by

$$\frac{1|}{|l_1 z} + \frac{1|}{|l_2} + \frac{1|}{|l_3 z} + \cdots \quad \text{and} \quad \frac{\lambda_1|}{|z-c_1} - \frac{\lambda_2|}{|z-c_2} - \frac{\lambda_3|}{|z-c_3} - \cdots,$$

where all l_{2i+1} are positive, all l_{2i} are negative, and λ_i and c_i are positive.

Stieltjes was not the first to discuss either the Moment-Problem, or the continued fraction (3). The first considerations along these lines are due to the great Russian mathematician Tchebycheff who in a series of papers started in 1855 discusses integrals of the type $\int_{-\infty}^{\infty} \frac{p(y)\, dy}{x - y}$, where $p(x)$ is non-negative in $(-\infty, \infty)$, and sums of the type $\sum_{-\infty}^{\infty} \frac{\theta_i^2}{x - x_i}$, $\theta_i \neq 0$ (both cases are now covered by a Stieltjes Integral). Tchebycheff's main tool is the theory of continued fractions which he uses with extreme ingenuity. However, Tchebycheff was not interested in the existence or construction of a solution of the Moment-Problem,

$$(6) \qquad \int_{-\infty}^{\infty} p(x)x^n\, dx = \mu_n, \qquad n = 0, 1, 2, \cdots,$$

but mainly in the following two problems: a) How far does a given sequence of moments determine the function $p(x)$? More particularly, given

$$\int_{-\infty}^{\infty} p(x)x^n\, dx = \int_{-\infty}^{\infty} e^{-x^2}x^n\, dx, \qquad n = 0, 1, 2, \cdots;$$

can we conclude that $p(x) = e^{-x^2}$, or, as we say now, that the distribution characterized by the function $\int_{-\infty}^{x} p(t)\, dt$ is a normal one? This is a fundamental problem in the theory of probability and in mathematical statistics. b) What are the properties of the polynomials $\omega_n(z)$, denominators of successive approximants to the continued fraction (3)? This opened a vast new field, the general theory of orthogonal polynomials, of which only the classical polynomials of Legendre, Jacobi, Abel—Laguerre and Laplace—Hermite were known before Tchebycheff. In the work of Tchebycheff we find numerous applications of orthogonal polynomials to interpolation, approximate quadratures, expansion of functions in series. Later they have been applied to the general theory of polynomials, theory of best approximations, theory of probability and mathematical statistics.

Another work which preceded that of Stieltjes is the classical work of Heine (1861, 1878, 1881). Here we find a brief discussion of the continued fraction associated with $\int_a^b \frac{p(y)\, dy}{x - y}$, where the given function $p(x)$ is non-negative in (a, b), and also an application of the orthogonal polynomials $\omega_n(x)$ to approximate quadratures.

One may venture the opinion that the use of this integral and of continued fractions was suggested to Stieltjes by the work of Tchebycheff and Heine. We must emphasize the importance of the new analytical tool, the Stieltjes Integral, which made it possible to treat the Problem of Moments in its most general form, namely,

$$(7) \qquad \int_{-\infty}^{\infty} x^n\, d\psi(x) = \mu_n, \qquad n = 0, 1, 2, \cdots.$$

One of the most talented pupils of Tchebycheff, A. Markoff, continued the work of his teacher applying it, in particular, to the theory of probability ("method of moments" applied to the proof of the fundamental limit-theorem), and to the closely related problem of finding precise bounds for $\int_c^d d\psi(x)$, $a < c < d < b$, where the function $\psi(x)$ is non-decreasing in (a, b), its first $n + 1$ moments being given. This important problem was proposed and its solution, based on the now celebrated "Tchebycheff inequalities", was given without proof by Tchebycheff in 1873. The proof was supplied by Markoff in his Thesis in 1884. It is interesting to note that Tchebycheff inequalities have been proved simultaneously and in the same manner by Stieltjes. Markoff further generalizes the Moment-Problem (1896) by requiring the solution $p(x)$ to be bounded:

$$\int_{-\infty}^{\infty} x^n p(x)\, dx = \mu_n, \qquad n = 0, 1, 2, \cdots, \quad \text{with} \quad 0 \leqq p(x) \leqq L.$$

In his investigations, as in those of his teacher, continued fractions play a predominant role.

As often happens in the history of science, the Problem of Moments lay dormant for more than 20 years. It revived again in the work of H. Hamburger, R. Nevanlinna, M. Riesz, T. Carleman, Hausdorff, and others.

An important approach to, and extension of, the work of Stieltjes to the whole real axis $(-\infty, \infty)$ was achieved by H. Hamburger (1920, 1921). This extension is by no means trivial. The consideration of negative values of x introduces new factors in the situation. Hamburger makes extensive use of Helly's theorem of choice. He fully discusses the convergence in the complex plane of both the associated and the corresponding (if it exists) continued fractions. He shows that a necessary and sufficient condition for the existence of a solution of the Moment-Problem (7) is the positiveness of all determinants Δ_n in (4), and also gives criteria for the Moment-Problem (7) to be determined or indeterminate. A curious fact is revealed, namely, that the Moment-Problem (7) may be indeterminate while the corresponding Stieltjes Moment-Problem (1), with the same μ_n, is determined.

R. Nevanlinna (1922) makes use of the modern theory of functions and exhibits the solutions of the Moment-Problem (7) and their properties in terms of the functions

$$I(z; \psi) = \int_{-\infty}^{\infty} \frac{d\psi(y)}{z - y}, \qquad z \text{ complex.}$$

To him is due the important notion of "extremal solutions".

About the same time (1921, 1922, 1923) M. Riesz solved the Moment-Problem (7) on the basis of "quasi-orthogonal polynomials", i.e. linear combinations $A_n\omega_n(z) + A_{n-1}\omega_{n-1}(z)$. He also showed the close connection between the Problem of Moments and the so-called "closure property" (Parseval Formula) for the orthogonal polynomials $\omega_n(z)$.

Carleman (1923, 1926) shows the connection between the Problem of Moments (7) and the theories of quasi-analytic functions and of quadratic forms in infinitely many variables (through the medium of the asymptotic series $\sum_0^\infty \mu_i z^{-i-1}$). To him is due the most general criterion, so far known, for the Moment-Problem to be determined.

Hausdorff (1923) gives criteria for the Moment-Problem (7) to possess a (necessarily unique) solution in a finite interval, that is, when $\psi(x)$ in (7) is required to remain constant outside a given finite interval. An effective construction of the solution is given and criteria are derived for the solution to have specified properties—continuity, differentiability, etc.

The interest in the Problem of Moments remains strong up to the present day. Among the most important contributions we may mention the work of Achyeser and Krein (1934). They have generalized the work of Markoff, making use of the tools of the theory of quadratic forms; they also extended the theory to the "trigonometric Moment-Problem"

$$\int_0^{2\pi} e^{inx}\, d\psi(x) = \mu_n, \qquad n = 0, 1, 2, \cdots. \tag{8}$$

Compare, in this connection, the work of S. Verblunsky (1932).

Carleman, and later, Stone developed a rather complete treatment of the Moment-Problem on the basis of the theory of Jacobi quadratic forms and singular integral equations and operators in Hilbert space. Finally, Haviland and Cramér extended M. Riesz' theory to the case of several dimensions.

Various generalizations have been made to the cases where the set of functions $\{x^n\}$ is replaced by a more general set $\{\varphi_n(x)\}$, or the integrals by more general linear operators in abstract spaces. These generalizations, however, will not be considered in the present monograph.

The discussion in the first two chapters follows the work of M. Riesz and R. Nevanlinna, in chapter III that of Markoff, Achyeser and Krein, and Hausdorff.

We shall now state explicitly, but mostly without proof, some fundamental facts which will be used in various places in the following chapters.

2. Distribution functions. Let $\mathfrak{R}_k$ be a k-dimensional Euclidean space. A function $\Phi(e)$ of sets e in $\mathfrak{R}_k$ is called a *distribution set-function* if it is non-negative, defined (and finite) over the family of all Borel sets in $\mathfrak{R}_k$, and is completely additive:

$$\sum_{i=1}^{\infty} \Phi(e_i) = \Phi\left(\sum_{i=1}^{\infty} e_i\right), \qquad e_i e_j = 0, \quad \text{if} \quad i \neq j.$$

The *spectrum* $\mathfrak{S}(\Phi)$ of a distribution set-function $\Phi(e)$ is defined as the set of all points $x \in \mathfrak{R}_k$, such that $\Phi(G) > 0$ for every open set G containing x.

The *point spectrum* of Φ is the set of all points x such that $\Phi((x)) > 0$.

By an interval $I \subset \mathfrak{R}_k$ we mean the set of points $x = (x_1, x_2, \cdots, x_k)$ whose coordinates satisfy conditions $a_i < x_i \leqq b_i$, $i = 1, 2, \cdots, k$, with obvious modifications in the case of an open or closed interval.

An interval I is an *interval of continuity* of the distribution set-function Φ (or more generally, for an additive function $\Phi(I)$ defined over all intervals) if, on introducing

$$I_\delta^{\mp}: \quad a_i \mp \delta < x_i \leqq b_i \pm \delta, \qquad i = 1, 2, \cdots, k,$$

we have

$$\Phi(I_\delta^{\mp}) \to \Phi(I), \qquad \text{as} \quad \delta \to 0.$$

Two distribution set-functions are said to be *substantially equal* if they have the same intervals of continuity and their values coincide over all such intervals.

Let $\Psi(I)$ be a non-negative set-function defined (and finite) over all intervals I in $\Re_k$ and satisfying the condition

$$\Psi(I) \leqq \sum_{i=1}^{n} \Psi(I_i), \quad \text{whenever} \quad I = \sum_{i=1}^{n} I_i, \qquad I_i I_j = 0 \quad \text{for} \quad i \neq j.$$

It is always possible to extend $\Psi(I)$ to a distribution set-function $\Phi(e)$ defined at least for all Borel sets having the same intervals of continuity as $\Psi(I)$, and coinciding with $\Psi(I)$ for such intervals.

A necessary and sufficient condition that two distribution set-functions Φ_1 and Φ_2 be substantially equal is that $\int_{\Re_k} f(t)\, d\Phi_1 = \int_{\Re_k} f(t)\, d\Phi_2$ for any continuous function $f(t)$ which vanishes for all sufficiently large values of $|t|$.

In the one-dimensional case a distribution set-function $\Phi(e)$ generates a point-function $\psi(t)$, which may be defined, for instance, by setting $\psi(t) = \Phi(I_t) + C$, where I_t is the infinite interval $-\infty < x \leqq t$, and C is an arbitrary constant. This point function is increasing and bounded in $(-\infty, \infty)$ and is determined uniquely at all its points of continuity, up to an additional constant. Conversely, every point function which is increasing and bounded generates a distribution set-function which is determined substantially uniquely.

For this reason any bounded increasing point-function may be called simply a distribution function.

Two distribution functions are said to be substantially equal if they have the same points of continuity and if their values at common points of continuity differ only by a constant. A function $\psi(t)$ which is increasing and bounded in a finite closed interval $[a, b]$ can be extended over the interval $(-\infty, \infty)$ by setting $\psi(t) = \psi(a)$, $t < a$, $\psi(t) = \psi(b)$, $t > b$. It then becomes a distribution function. Two functions $\psi_1(t)$, $\psi_2(t)$ which are increasing and bounded over a finite closed interval $[a, b]$ are said to be substantially equal if they have the same interior points of continuity and if their values at these points, and also at the end-points $t = a$, $t = b$, differ by a constant. Analogous considerations hold, of course, in the general k-dimensional case.

For proofs we refer to [Bochner, 1; Haviland, 2].

3. Theorems of Helly. A sequence of additive functions of intervals $\{\Psi_n(I)\}$ is said to *converge substantially* to a function of intervals $\Psi(I)$ if $\lim_{n\to\infty} \Psi_n(I) = \Psi(I)$ for all (finite) intervals of continuity of Ψ.

"FIRST THEOREM OF HELLY". *Given a sequence* $\{\Psi_n(I)\}$ *of non-negative additive and uniformly bounded functions of intervals, then there exists a subsequence* $\{\Psi_{n_\nu}(I)\}$ *and a distribution function* Φ *to which this subsequence converges substantially. Furthermore, if the sequence* $\{\Psi_n\}$ *itself does not converge substantially to* Φ, *then there exists another subsequence* $\{\Psi_{n'_\nu}(I)\}$ *converging substantially to another distribution function* Φ' *which is not substantially equal to* Φ.

"SECOND THEOREM OF HELLY". *Given a sequence* $\{\Psi_n(I)\}$ *of non-negative additive and uniformly bounded functions of intervals, which converges substantially to a distribution function* Φ. *Then*

$$\lim_n \int_{\mathfrak{R}_k} f(t)\, d\Psi_n = \int_{\mathfrak{R}_k} f(t)\, d\Phi$$

for any function $f(t)$ *continuous in* $\mathfrak{R}_k$ *and such that, as* $I_N \uparrow \mathfrak{R}_k$, $\int_{I_N} f(t)\, d\Psi_n \to \int_{\mathfrak{R}_k} f(t)\, d\Psi_n$ *uniformly in* n.

In the one-dimensional case this theorem may be easily restated in terms of sequences of uniformly bounded increasing point-functions, instead of functions of intervals. For the proof see [Bochner, 1].

4. Extension theorem for non-negative functionals. *Let* $\mathfrak{M}$ *be a linear manifold* of real-valued functions* $x(t)$ *defined on any abstract space* Ω, $t \in \Omega$. *Let* $\mathfrak{M}_0$ *be a linear sub-manifold of* $\mathfrak{M}$ *and let* $f_0(x)$ *be a* (Ω_0) *non-negative additive and homogeneous functional defined on* $\mathfrak{M}_0$, *that is*

$$f_0(x_1 + x_2) = f_0(x_1) + f_0(x_2), \qquad x_1, x_2 \in \mathfrak{M}_0,$$

$$f_0(cx) = cf(x), \qquad x \in \mathfrak{M}_0,$$

$$f_0(x) \geqq 0, \quad \textit{whenever} \quad x(t) \geqq 0 \quad \textit{for all} \quad t \in \Omega_0 \subset \Omega.^{**}$$

This functional $f_0(x)$ *can be extended to an additive, homogeneous and* (Ω_0) *non-negative functional* $f(x)$ *defined on the whole manifold* $\mathfrak{M}$ *so that* $f(x)$ *coincides with* $f_0(x)$ *on* $\mathfrak{M}_0$.

Assume $\mathfrak{M}_0 \subset \mathfrak{M}$ and $y_1 \in \mathfrak{M} - \mathfrak{M}_0$. Consider the linear manifold $\mathfrak{M}_1$ determined by $\mathfrak{M}_0$ and y_1. The elements x_1 of $\mathfrak{M}_1$ admit of a unique representation $x_1 = x_0 + ty_1$, where x_0 is any element of $\mathfrak{M}_0$ and t is any real number. Introduce the functional $f(x_1)$ defined on $\mathfrak{M}_1$ by

$$f(x_1) = f(x_0 + ty_1) = f_0(x_0) + tr_1, \qquad r_1 = f(y_1).$$

It is clear that this functional is additive and homogeneous, and that it coincides with $f_0(x)$ when $x \in \mathfrak{M}_0$. It remains to determine r_1 so that f will be (Ω_0) non-negative. Take any x_0 such that $x_0 - y_1 \geqq 0$ on Ω_0. Then the condition $f(x_0 - y_1) \geqq 0$ implies $f_0(x_0) - r_1 \geqq 0$. Hence, on setting $M = \inf f_0(x_0)$, if

* That is, a set of functions $x(t)$ which contains $cx(t)$, $x(t) + y(t)$, whenever $x(t)$, $y(t)$ belong to the set, and c is a real constant.

** Ω_0 is any given set in Ω; in particular, Ω_0 may coincide with Ω.

there exist elements $x_0 \geqq y_1$, on Ω_0 and $M = \infty$ in the opposite case, we have $r_1 \leqq M$. In the same way, on setting $m = \sup f_0(x_0)$, if there exist elements $x_0 \leqq y_1$ on Ω_0, and $m = -\infty$ in the opposite case, we find $m \leqq r_1$.

Retracing our steps we readily find that if we take for $r_1 = f(y_1)$ any number satisfying the condition $m \leqq r_1 \leqq M$, the functional $f(x_1)$ defined above will be (Ω_0) non-negative. Thus $f_0(x)$ is extended to the linear manifold $\mathfrak{M}_1$. The extension to the whole linear manifold $\mathfrak{M}$ can now be performed by the method of transfinite induction.

The proof above proceeds along the same lines as a proof by Kantorovich [1]. Compare also Haviland [4, 5].

5. Stieltjes inversion formula. Let $\psi(t)$ be any function of bounded variation on $(-\infty, \infty)$. The integral $I(z) \equiv I(z; \psi) = \int_{-\infty}^{\infty} \frac{d\psi(t)}{z - t}$ is an analytic function of z in the upper and in the lower half-planes, its values being conjugate at two conjugate points. The function $\psi(t)$ can be expressed in terms of $I(z)$ by the following formula:

$$\tfrac{1}{2}[\psi(t_1 + 0) + \psi(t_1 - 0)] - \tfrac{1}{2}[\psi(t_0 + 0) + \psi(t_0 - 0)] = \lim_{\epsilon \to +0} -\frac{1}{2\pi i}\int_{t_0}^{t_1} [I(t + i\epsilon) - I(t - i\epsilon)]\, dt.$$

(Cf. Stone, [1]). Thus, $\psi(t)$ is substantially uniquely determined by $I(z; \psi)$.

TABLE OF CONTENTS

CHAPTER I

A GENERAL THEORY OF THE PROBLEM OF MOMENTS.

1. Let $\mathfrak{R}$ be a k-dimensional Euclidean space. Let there be given an infinite multiple sequence of real constants

$$\mu_{j_1 j_2 \cdots j_k}; \qquad j_1, j_2, \cdots, j_k = 0, 1, 2, \cdots.$$

We are interested in finding necessary and sufficient conditions that there shall exist a k-dimensional distribution function Φ whose spectrum $\mathfrak{S}(\Phi)$ is to be contained in a closed set $\mathfrak{S}_0$, given in advance, and which is a solution of the "problem of moments" [Haviland, 4, 5]

$$(1.1) \qquad \mu_{j_1 \cdots j_k} = \int_{\mathfrak{R}} t_1^{j_1} \cdots t_k^{j_k} \, d\Phi, \qquad j_1, j_2, \cdots, j_k = 0, 1, 2, \cdots.$$

To abbreviate we call this problem simply the ($\mathfrak{S}_0$) moment problem. We say that the moment problem is determined if its solution is substantially unique; otherwise we call it indeterminate.

To simplify we shall discuss only the two-dimensional case, $k = 2$. There is no difficulty in extending the results to the case of any number of dimensions.

Let $P(u, v)$ be any polynomial in u, v in $\mathfrak{R}$,

$$P(u, v) = \sum_{i,j} x_i y_j u^i v^j,$$

where x_i, y_j are real- or complex-valued constants. Introduce the functional $\mu(P)$ defined by

$$\mu(P) \equiv \sum_{i,j} \mu_{ij} x_i y_j.$$

In particular,

$$\mu(u^i v^j) = \mu_{ij}.$$

THEOREM 1.1. *A necessary and sufficient condition that the* ($\mathfrak{S}_0$) *moment problem defined by the sequence of moments* $\{\mu_{ij}\}$ *shall have a solution is that the functional* $\mu(P)$ *be* ($\mathfrak{S}_0$) *non-negative, that is*

$$(1.2) \qquad \mu(P) \geqq 0, \quad \textit{whenever} \quad P(u, v) \geqq 0 \quad \textit{on} \quad \mathfrak{S}_0.$$

This theorem is an immediate application of the theorem on the extension of non-negative functionals (Introduction, **4**). Let $\mathfrak{M}$ be the linear manifold of all single-valued functions $y = y(u, v)$ which admit of an estimate

$$(1.3) \qquad |y(u, v)| \leqq A(u^{2r} + v^{2r}) + B,$$

where A, B are non-negative constants and r is a non-negative integer. Let $\mathfrak{M}_0$ be the linear sub-manifold of $\mathfrak{M}$, consisting of all polynomials P. It is clear that all functions $\eta = A(u^{2r} + v^{2r}) + B$ are in $\mathfrak{M}_0$.

Now if our $(\mathfrak{S}_0)$ moment-problem has a solution Φ, then whenever $P \geqq 0$ on $\mathfrak{S}_0$, we obviously have

$$\mu(P) = \int_{\mathfrak{R}} P(u, v)\, d\Phi = \int_{\mathfrak{S}_0} P(u, v)\, d\Phi \geqq 0.$$

Thus the condition (1.2) of Theorem 1.1 is necessary. To prove its sufficiency, suppose (1.2) is satisfied. Then $\mu(P)$ appears as a homogeneous additive $(\mathfrak{S}_0)$ non-negative functional defined on $\mathfrak{M}_0$. By Introduction, **4**, this functional can be extended over the whole manifold $\mathfrak{M}$, with preservation of all these properties. In particular, we may define $\mu(y_I)$, where y_I is the characteristic function of any two-dimensional interval I, since clearly $y_I \in \mathfrak{M}$. Thus we obtain a function $\psi(I) = \mu(y_I)$ of intervals, which possesses the properties

(i) $$\psi(I) \geqq 0,$$

since $y_I \geqq 0$ and μ is non-negative;

(ii) whenever $I = \sum_{i=1}^{n} I_i, \quad I_i I_j = 0$ for $i \neq j$, then

$$\psi(I) = \sum_{i=1}^{n} \psi(I_i),$$

since μ is additive and $y_I = \sum_{i=1}^{n} y_{I_i}$;

(iii) $\psi(I)$ is bounded,

since

$$\psi(I) \leqq \psi(\mathfrak{R}) = \mu(y_{\mathfrak{R}}) = \mu(1) = \mu_{00}.$$

The conditions of Introduction, **2**, are thus satisfied, and we can construct the associated distribution function Φ which is substantially equal to $\psi(I)$. This function Φ is a solution of our $(\mathfrak{S}_0)$ moment problem. To prove this we have only to establish that

(1.4) $$\mathfrak{S}(\Phi) \leqq \mathfrak{S}_0,$$

(1.5) $$\int_{\mathfrak{R}} u^i v^j\, d\Phi = \mu_{ij}, \qquad i, j = 0, 1, 2, \cdots.$$

To prove (1.4) it suffices to show that $(u_0, v_0) \in \mathfrak{R} - \mathfrak{S}_0$ implies $(u_0, v_0) \in \mathfrak{R} - \mathfrak{S}(\Phi)$. Let $(u_0, v_0) \in \mathfrak{R} - \mathfrak{S}_0$ and let $I_0 < \mathfrak{R} - \mathfrak{S}_0$ be a common interval of continuity of Φ and ψ containing (u_0, v_0) in its interior. Since $y_{I_0} = 0$ on $\mathfrak{S}_0$ we may write $y_{I_0} \leqq 0$, $y_{I_0} \geqq 0$ on $\mathfrak{S}_0$, whence $\mu(y_{I_0}) \leqq 0$, $\mu(y_{I_0}) \geqq 0$, and $\mu(y_{I_0}) = 0$. Thus $\psi(I_0) = \mu(y_{I_0}) = 0$, which implies $\Phi(I_0) = 0$; hence $(u_0, v_0) \in \mathfrak{R} - \mathfrak{S}(\Phi)$.

It remains to prove (1.5). Let ϵ, ϵ_1 be two given positive numbers. Let I_0 be again a common interval of continuity of Φ and ψ, so large that for a suitable choice of r

$$| u^i v^j | < \epsilon(u^{2r} + v^{2r}) \quad \text{on} \quad \mathfrak{R} - I_0 .$$

The integers i, j, and r will be now fixed. Let $I_1, I_2, \cdots, I_n$ be a finite sequence of common intervals of continuity of Φ and ψ, disjoint and such that

$$I_0 = I_1 + I_2 + \cdots + I_n ,$$

while the oscillation of $u^i v^j$ on each $\bar{I}_\nu$, $\nu = 1, 2, 3, \cdots, n$ is less than ϵ_1. In each I_ν select a point (u_ν, v_ν) and introduce a simple function

$$y_0(u, v) = \begin{cases} u_\nu^i v_\nu^j \quad \text{on} \quad I_\nu , & \nu = 1, 2, \cdots, n, \\ 0 \text{ elsewhere.} & \end{cases}$$

It is clear that

$$y_0(u, v) = \sum_{\nu=0}^{n} u_\nu^i v_\nu^j y_{I_\nu} .$$

Since

$$y_0(u, v) - \epsilon_1 < u^i v^j < y_0(u, v) + \epsilon_1 \quad \text{on} \quad I_\nu , \qquad \nu = 1, 2, \cdots, n,$$

while

$$-\epsilon(u^{2r} + v^{2r}) < u^i v^j < \epsilon(u^{2r} + v^{2r}) \quad \text{on} \quad \mathfrak{R} - I_0 ,$$

we have everywhere in $\mathfrak{R}$

$$y_0(u, v) - \epsilon(u^{2r} + v^{2r}) - \epsilon_1 < u^i v^j < y_0(u, v) + \epsilon(u^{2r} + v^{2r}) + \epsilon_1 .$$

In view of the $(\mathfrak{S}_0)$ non-negativeness of the functional μ, we have

$$\mu[y_0(u, v) - \epsilon(u^{2r} + v^{2r}) - \epsilon_1] \leqq \mu(u^i v^j) \leqq \mu[y_0(u, v) + \epsilon(u^{2r} + v^{2r}) + \epsilon_1],$$

$$\mu(y_0) - \epsilon(\mu_{2r,0} + \mu_{0,2r}) - \epsilon_1\mu_{00} \leqq \mu_{ij} \leqq \mu(y_0) + \epsilon(\mu_{2r,0} + \mu_{0,2r}) + \epsilon_1\mu_{00} .$$

But

$$\mu(y_0) = \sum_{\nu=1}^{n} u_\nu^i v_\nu^j \mu(y_{I_\nu}) = \sum_{\nu=1}^{n} u_\nu^i v_\nu^j \psi(I_\nu) = \sum_{\nu=1}^{n} u_\nu^i v_\nu^j \Phi(I_\nu).$$

If we allow here $\epsilon_1 \to 0$ and $\max | I_\nu | \to 0$, we see that

$$\int_{I_0} u^i v^j \, d\Phi - \epsilon(\mu_{2r,0} + \mu_{0,2r}) \leqq \mu_{ij} \leqq \int_{I_0} u^i v^j \, d\Phi + \epsilon(\mu_{2r,0} + \mu_{0,2r}).$$

On allowing here $\epsilon \to 0$ (and $I_0 \to \mathfrak{R}$) it readily follows that $u^i v^j$ is absolutely integrable over $\mathfrak{R}$ with respect to Φ and that

$$\mu_{ij} = \int_{\mathfrak{R}} u^i v^j \, d\Phi, \qquad i, j = 0, 1, 2, \cdots .$$

2. Theorem 1.1 can be readily applied to derive necessary and sufficient conditions for the existence of solutions of various specialized moment problems characterized by a special choice of $\mathfrak{S}_0$.

(a) HAMBURGER MOMENT PROBLEM. Here $\mathfrak{S}_0$ coincides with the axis of reals. Hence $\mu_{ij} = 0$ for $j \geqq 1$, so that we have a simple sequence of moments

$$\mu_n \equiv \mu_{n0}, \qquad n = 0, 1, 2, \cdots,$$

and the problem reduces to that of determining a one-dimensional distribution function $\psi(u)$ such that

$$(1.6) \qquad \mu_n = \int_{-\infty}^{\infty} u^n \, d\psi(u), \qquad n = 0, 1, 2, \cdots.$$

Thus it suffices to consider polynomials and functions of u alone, and to define the functional μ by

$$(1.7) \qquad \mu(P_n) = \sum_{j=0}^{n} \mu_j x_j, \qquad P_n(u) = \sum_{j=0}^{n} x_j u^j.$$

Theorem 1.1 states now that a necessary and sufficient condition for the existence of a solution of (1.6) is that $\mu(P) \geqq 0$, whenever $P(u) \geqq 0$ for all real values of u. If we take for $P(u)$ the particular polynomial $P(u) = (x_0 + x_1 u + \cdots + x_n u^n)^2$, x_i real, we have

$$(1.8) \qquad \mu(P) = \sum_{i,j=0}^{n} \mu_{i+j} x_i x_j \equiv Q_n(x)$$

(Hankel quadratic form). Thus *a necessary condition for the existence of a solution of* (1.6) *is that the quadratic forms* $Q_n(x)$, $n = 0, 1, 2, \cdots$, *be non-negative. This condition is also sufficient.*

In fact [Pólya und Szegö, 1, Vol. II, p. 82] any polynomial $P(u) \geqq 0$ for all real u can be represented by

$$P(u) = p_1(u)^2 + p_2(u)^2,$$

where $p_1(u)$, $p_2(u)$ are polynomials with real coefficients, whence

$$u(P) = \mu(p_1^2) + \mu(p_2^2) \geqq 0,$$

if $Q_n(x) \geqq 0, n = 0, 1, 2, \cdots$.

Let $\psi(u)$ be a solution of (1.6). Since

$$Q_n(x) = \mu(p_n^2) = \int_{-\infty}^{\infty} p_n^2(t) \, d\psi,$$

it is clear that if $\mathfrak{S}(\psi)$ is not reducible to a finite set of points, we always have

$$(1.9) \qquad Q_n(x) = \sum_{i,j=0}^{n} \mu_{i+j} x_i x_j > 0, \qquad n = 0, 1, 2, \cdots,$$

provided not all $x_0, x_1, \cdots, x_n$ are zero, which will be assumed in what follows. From the theory of quadratic forms it is well known that conditions (1.9) are equivalent to

$$\Delta_n = | \mu_{i+j} |_{i,j=0}^{n} > 0, \qquad n = 0, 1, 2, \cdots.$$

On the other hand, if there exists a solution $\psi(u)$ whose spectrum consists precisely of $(k + 1)$ distinct points (we shall see later [**3**, Corollary 1.1] that the moment problem is then determined), $t_1, t_2, \cdots, t_{k+1}$, then it is readily seen that for each $n \geqq k + 1$, $Q_n(x) = 0$ for a suitable choice of $x_0, x_1, \cdots, x_n$, which implies $\Delta_n = 0$, $n = k + 1, k + 2, \cdots$, while $\Delta_0 > 0, \cdots, \Delta_k > 0$. It can be proved, conversely, that if these conditions are satisfied, then there exists a uniquely determined solution of the moment problem with the property mentioned above [Fischer, 1; Achyeser and Krein, 1, 6].

All these results can be stated in

THEOREM 1.2. *In order that a Hamburger moment problem*

$$\mu_n = \int_{-\infty}^{\infty} t^n \, d\psi, \qquad n = 0, 1, 2, \cdots, \tag{1.6}$$

shall have a solution it is necessary that

$$\Delta_n = | \mu_{i+j} |_{i,j=0}^{n} \geqq 0, \qquad n = 0, 1, 2, \cdots. \tag{1.10}$$

In order that there exist a solution whose spectrum is not reducible to a finite set of points it is necessary and sufficient that

$$\Delta_n > 0, \qquad n = 0, 1, 2, \cdots. \tag{1.11}$$

In order that there exist a solution whose spectrum consists of precisely $(k + 1)$ *distinct points it is necessary and sufficient that*

$$\Delta_0 > 0, \cdots, \Delta_k > 0, \qquad \Delta_{k+1} = \Delta_{k+2} = \cdots = 0. \tag{1.12}$$

The moment problem is determined in this case.

(b) STIELTJES MOMENT PROBLEM. In this case $\mathfrak{S}_0$ coincides with the positive part of the axis of reals, $u \geqq 0$. As in the preceding case, we have to consider only moments $\mu_n = \mu_{n0}$, and only polynomials and functions of a single variable. The moment problem reduces to

$$\mu_n = \int_0^{\infty} t^n d\psi, \qquad n = 0, 1, 2, \cdots,$$

and a necessary and sufficient condition for the existence of a solution is that

$$\mu(P) = \sum_{j=0}^{n} \mu_j x_j \geqq 0,$$

whenever

$$P(u) = x_0 + x_1 u + \cdots + x_n u^n \geqq 0 \quad \text{for} \quad u \geqq 0.$$

An application of this condition to the two special polynomials $(x_0 + x_1 u + \cdots + x_n u^n)^2$, $u(x_0 + x_1 u + \cdots + x_n u^n)^2$ yields at once

$$\left.\begin{aligned} Q_n(x) &\equiv \sum_{i,j=0}^{n} \mu_{i+j} x_i x_j \geqq 0, \\ Q_n^{(1)}(x) &\equiv \sum_{i,j=0}^{n} \mu_{i+j+1} x_i x_j \geqq 0, \end{aligned}\right\} \qquad n = 0, 1, 2, \cdots,$$